



Série d'activités d'Actua sur l'IA

Activité 7

Les IA à commande vocale : entraîner un modèle de reconnaissance de la parole

Les IA à commande vocale : entraîner un modèle de reconnaissance de la parole

Vous avez accédé directement à cette activité? Saviez-vous que notre série en compte huit autres? Consultez notre site web pour en connaître les détails et l'ordre recommandé pour les réaliser. Elles présentent toutes des notions d'intelligence artificielle ainsi qu'un vocabulaire propre au domaine. [Un glossaire](#) permettant de vérifier le sens des mots nouveaux ou inconnus les accompagne. Amusez-vous tout en réparant une station spatiale!

Récapitulons : toi et ton groupe faites partie d'une équipe d'astronautes et de scientifiques en mission à bord de la station spatiale Actua Comble de malchance, la station vient d'être bombardée par des rayons magnétiques. Tous vos appareils électroniques tombent en panne les uns après les autres. Le seul qui peut vous tirer d'affaire, c'est DANN, le réseau neuronal spécialisé d'Actua, mais il a un peu perdu la carte depuis l'incident. Révisez vos connaissances techniques et apprenez quelques notions d'intelligence artificielle pour sauver l'équipage!*

Grâce à vos connaissances, vous avez réussi à réinitialiser le processeur audio de DANN dans l'activité [«Ce que voient les ordinateurs : s'initier à la vision artificielle»](#), mais lorsque vous lui avez demandé d'ouvrir une porte, il vous a préparé un lait frappé. Tout indique que son processeur audio a été endommagé davantage que prévu. Son système de reconnaissance de la parole détecte votre voix, mais il ne comprend pas ce qu'il entend. Vous allez devoir le réentraîner pour pouvoir lui donner des commandes. Quand tout fonctionnera bien, vous pourrez vous passer à l'étude du code d'éthique de DANN dans l'activité [«Éthique de l'IA : Il faut empêcher DANN de devenir malveillant»](#)!

**Pour «Dedicated Actua Neural System»*



Présentation de l'activité

Dans cette activité, les élèves entraîneront un logiciel d'écoute machine à exécuter une tâche de classification. Ils exploreront d'abord un classifieur audio préentraîné, puis ils utiliseront Teachable Machine dans le but d'entraîner leur propre modèle de classification des sons à reconnaître certains mots-clés.

Activité conçue par Actua, en 2022.

Contexte de l'activité	Durée	Public cible
Salle de classe avec accès Internet	60 minutes	Niveaux 9-12 (13-18 ans)

Résultats d'apprentissage

À la suite de cette activité, les participants pourront :

- Décrire et expliquer en quoi consiste la *classification*, une des applications principales de l'intelligence artificielle.
- Créer et entraîner un modèle d'intelligence artificielle capable d'exécuter des tâches de classification à l'aide de Teachable Machine de Google.
- Interpréter les données de sortie d'un modèle (étiquette et niveau de confiance) en termes généraux (très élevé, élevé, peu élevé).
- Évaluer la pertinence d'un modèle d'intelligence artificielle entraîné à exécuter une tâche donnée.

OUTILS	COMPÉTENCES	ATTITUDES
Connaissances, ressources et expériences	Compétences numériques, compétences en STIM, employabilité et	Intelligence numérique, action communautaire et pensée



OUTILS	COMPÉTENCES	ATTITUDES
<ul style="list-style-type: none"> • Décrire et expliquer en quoi consiste la classification. • Interpréter les données rapportées par un modèle d'apprentissage automatique. • Élaborer un modèle d'apprentissage automatique. 	<p>aptitudes essentielles à la vie</p> <ul style="list-style-type: none"> • Analyse critique • Évaluation statistique • Programmation 	<p>computacionnelle</p> <ul style="list-style-type: none"> • Évaluer les avantages des logiciels d'apprentissage automatique en fonction de tâches variées. • Explorer les possibilités qu'offre le système d'entrée et de sortie d'un logiciel et leurs applications.

Logistique (durée, taille du groupe, matériel)

Titre de la section	Durée	Taille du groupe	Matériel
<p>Introduction : Explorer un modèle préentraîné</p>	<p>10 min</p>	<p>Toute la classe</p>	<p>Pour chaque petit groupe :</p> <ul style="list-style-type: none"> • Ordinateur avec microphone • Accès Internet :



Titre de la section	Durée	Taille du groupe	Matériel
			<ul style="list-style-type: none"> • p5 interactive sketch [modèle interactif] : https://editor.p5js.org/Actua/full/sYZkhDFgP • Stylo/crayon • Papier
Activité n° 1 : Définir un modèle de classification des sons	30 min	Petits groupes ou toute la classe	Pour chaque petit groupe : <ul style="list-style-type: none"> • Ordinateur avec microphone • Annexe A : Commandes RCS • Accès Internet : • Teachable Machine: https://teachablemachine.withgoogle.com/train/audio • p5 interactive sketch (en anglais): https://editor.p5js.org/Actua/full/p-UhfjBo_
Réflexion et récapitulation	5 min	Toute la classe	

Consignes de sécurité

Les consignes de sécurité ci-dessous ne sont pas exhaustives. Veillez à passer en revue l'activité et à inspecter l'environnement où elle sera réalisée afin de déterminer si des mesures additionnelles sont requises pour assurer la sécurité des élèves.



Sécurité en ligne

Certains volets de cette activité nécessitent l'usage d'appareils connectés à Internet.

- Examinez au préalable les vidéos, les sites web et le matériel prévus afin de vous assurer qu'ils conviennent à vos élèves.
- Au besoin, rappelez aux jeunes de se concentrer sur la tâche à faire et d'utiliser uniquement les liens fournis pour l'activité.

Marche à suivre

Le modèle de reconnaissance de la parole de DANN doit être complètement réinitialisé et entraîné. Toutefois, pour arriver à le faire, vous devrez rediriger la principale batterie d'antennes de la station vers le satellite de récupération (SR). À partir de vos connaissances sur le modèle de commande du processeur visuel, vous devrez entraîner un modèle simple d'écoute machine de façon à ce qu'il puisse exécuter les commandes du système de contrôle par jet de gaz. Ce système vous permettra de réaligner la batterie d'antennes vers l'un des satellites de réserve.

Introduction : Entraîner un modèle préentraîné

Tout d'abord, vous allez explorer un modèle de reconnaissance de la parole qui a été préentraîné, à l'image de ce que vous avez fait dans l'activité [«Ce que voient les ordinateurs»](#). Appelé «SpeechCommands18w», ce modèle a appris à reconnaître 18 mots en anglais :

- les chiffres «zero» à «nine».
- quatre directions : «up», «down», «left» et «right»
- «go» et «stop»
- «yes» et «no»

Également, le modèle peut indiquer s'il a détecté un mot inconnu («unknown word») ou s'il n'entend qu'un bruit de fond («background noise»). Il a été chargé dans un programme d'évaluation pour que vous le testiez :

<https://editor.p5js.org/Actua/full/sYZkhDFqP>.



1. Ouvrez le modèle interactif. Comme il s'agit d'un modèle de reconnaissance de la parole, il pourrait vous demander la permission d'utiliser le microphone de votre ordinateur.
2. Choisissez cinq mots dans le vocabulaire du modèle (c.-à-d. cinq sur les 18 mots qu'il a appris à reconnaître).
3. Dicter les mots que vous avez choisis en laissant 1 à 2 secondes de silence entre chacun. Prenez note du résultat : le modèle a-t-il reconnu le mot que vous avez prononcé?
4. Répétez la troisième étape 1 à 3 fois.
5. À partir des résultats obtenus aux étapes 3 et 4, répondez aux questions suivantes :
 - a. Y a-t-il des mots que le modèle a eu du mal à reconnaître? Lesquels?
 - b. A-t-il fait des erreurs?
 - c. À votre avis, quelles conditions aideraient le modèle à reconnaître les mots correctement?

Le modèle SpeechCommands18w sait reconnaître plusieurs mots qui peuvent être utiles, mais il n'est pas entraîné à reconnaître les commandes dont vous aurez besoin pour contrôler la station. Cela signifie que vous devrez entraîner un modèle capable de le faire.

Activité n° 1 : Définir un modèle de classification des sons

La marche à suivre pour entraîner un modèle de classification des sons à l'aide de Teachable Machine est semblable à celle qu'on utilise pour la classification d'images.

Étape n°1. Chargez l'outil Teachable Machine

1. L'outil Teachable Machine est accessible à l'adresse suivante : <https://teachablemachine.withgoogle.com/train/audio>.



2. Ce lien vous conduira au modèle de classification des sons de Teachable Machine. Vous avez probablement déjà utilisé son classifieur d'images au moment de créer votre modèle pour le processeur visuel de DANN.

Étape n° 2. Définissez des classes

1. Le processeur audio utilisera les noms des classes de votre modèle pour transmettre les commandes correspondantes.
2. À votre avis, quelles classes faudrait-il définir? Consultez l'annexe A pour connaître les fonctions du système de contrôle par jet de gaz (SCJG).
 - a. *Conseil : Le Contrôle de mission vous conseille de définir une classe pour chaque direction ainsi que pour le mot «stop». Teachable Machine a déjà créé la classe «background noise», mais ce pourrait être une bonne idée d'ajouter une classe «unknown» au modèle.*

Étape n° 3. Créez des données d'apprentissage

1. Vous vous servirez du microphone de votre ordinateur pour créer des données d'apprentissage. Teachable Machine enregistrera une plage sonore continue (fixée par défaut à 2 secondes), qu'il décomposera automatiquement en clips d'une seconde pour s'entraîner. Chaque classe doit comprendre un minimum de 8 échantillons sonores différents (hormis la classe «background noise», qui devra en inclure 20). N'hésitez pas à en enregistrer davantage si vous avez le temps, car plus le nombre d'échantillons sera grand, meilleure sera la reconnaissance!

Étape n° 4. Entraînez votre modèle

Vous devez maintenant entraîner votre modèle de reconnaissance de la parole à l'aide de vos données audio afin qu'il puisse ensuite reconnaître les commandes voulues, comme vous l'aviez fait avec votre modèle de classification d'images. Voici comment faire :

1. Dans la boîte étiquetée «Training», cliquez sur le bouton «Train Model».
2. Attendez que l'apprentissage se termine. Au bout d'un moment, vous devriez apercevoir un compteur ainsi qu'un chiffre sur 50. Le processus



est achevé lorsque le chiffre atteint 50 sur 50, mais il peut prendre quelques secondes.

3. Une fois l'entraînement terminé, la boîte « Preview » devrait contenir une section « Output » montrant en temps réel la classification des données audio réalisée par le modèle et le degré de confiance qu'il lui accorde.

Étape n° 5 : Testez le modèle

1. Le temps est venu de vérifier si votre modèle a été suffisamment bien entraîné pour pouvoir réaligner la batterie d'antennes de la station. Il devrait être capable de reconnaître avec précision les mots que vous avez choisis et de détecter si nécessaire le bruit de fond ou les mots inconnus.
2. Vous pouvez adapter les questions qui ont servi à l'évaluation du modèle de classification d'images pour vérifier si votre modèle fonctionne bien :
 - a. *Votre modèle reconnaît-il avec exactitude vos commandes lorsqu'elles sont dictées par vous ou d'autres membres du groupe ayant participé à la collecte des données d'apprentissage?*
 - b. *Votre modèle reconnaît-il avec exactitude vos commandes lorsqu'elles sont dictées par des personnes qui n'ont pas participé à la collecte?*
 - c. *Votre modèle reconnaît-il avec exactitude vos commandes lorsqu'elles sont dictées dans un cadre autre que celui où s'est déroulée la collecte? (par ex. dans une autre partie de la pièce, avec des bruits de fond ou niveaux sonores différents)?*
3. Si vous avez répondu « oui » à toutes ces questions, votre modèle est prêt à être utilisé. Dans le cas contraire, consultez la rubrique dépannage :
 - a. Y a-t-il dans vos données des sons qui ne sont pas de bonnes représentations de la classe à laquelle ils appartiennent?



Possibilités d'adaptation

Il est possible d'adapter différents aspects de cette activité (durée, environnement, matériel, taille du groupe ou instructions) pour la rendre plus accessible ou plus complexe. Les **modifications** ci-dessous vous permettront de diminuer le niveau de difficulté de l'activité et les **ajouts**, d'augmenter sa durée ou son niveau de difficulté.

Ajouts

- La section « Model and data provenance » fournit de l'information sur la création et l'entraînement du modèle SpeechCommands18w : <https://editor.p5js.org/Actua/full/sYZkhDFgP>. Après l'avoir lue, répondez aux questions suivantes :
 - Pourquoi s'agit-il d'information importante?
 - Comment pourrait-elle nous aider à construire de meilleurs modèles?
- Pour pousser l'exercice plus loin, créez de nouvelles classes de commandes vocales. Imaginez que vous donnez des commandes à une IA chargée d'opérer une station spatiale. Pensez aux commandes qu'il faudrait lui donner. Créez de nouvelles classes et entraînez votre modèle en veillant à produire la même quantité de données que dans l'exercice précédent. Testez-le ensuite. Lorsque vous augmentez le nombre de classes, le modèle réussit-il mieux à reconnaître les commandes, ou moins bien? Comment faire pour l'améliorer?

Modifications

- Pour épargner du temps, l'exploration du modèle proposée dans l'introduction peut être faite avec le groupe entier plutôt qu'en petits groupes. Demandez aux élèves de dicter à tour de rôle les mots du vocabulaire.
- Pour rendre l'activité plus facile pour de jeunes élèves : il est possible de préenregistrer un jeu de données dans Teachable Machine et de le



télécharger ensuite dans Google Drive. Les élèves peuvent ainsi travailler avec un jeu de données préexistantes, ce qui permet de rendre l'expérience moins complexe et d'obtenir de meilleurs résultats.

Références et remerciements

Cette activité a été réalisée grâce à l'outil Teachable Machine de Google.
<https://teachablemachine.withgoogle.com/train/audio>.

Plusieurs exercices ont été conçus à l'aide de p5.js, une bibliothèque JavaScript en ligne accessible à : <https://p5js.org/>.



Conditions d'utilisation

Avant de réaliser cette activité en tout ou en partie, vous reconnaissez et acceptez ce qui suit :

- Il vous appartient de passer en revue toutes les sections du présent document et la documentation connexe ainsi que d'appliquer les consignes de sécurité nécessaires à la protection de toutes les personnes concernées;
- Les mesures précisées à la rubrique « Consignes de sécurité » du présent document ne sont pas exhaustives ni ne remplacent votre propre cadre d'examen de la sécurité;
- Actua n'est pas responsable des dommages attribuables à l'usage du présent contenu;
- Vous pouvez adapter ce document à vos besoins (le remanier, le transformer ou créer du matériel à partir de celui-ci), à condition d'indiquer qu'Actua en est l'auteur original et que vous y avez apporté des changements. Ce contenu ne peut être transmis à de tierces parties sans la permission écrite d'Actua.

À propos d'Actua

Représentant plus de 40 universités et collèges à travers le pays, Actua est le principal réseau de sensibilisation des jeunes aux sciences, à la technologie, à l'ingénierie et aux mathématiques (STIM) au Canada. Chaque année, 350 000 jeunes prennent part à des ateliers pratiques, à des camps et à des projets communautaires inspirants dans plus de 500 localités d'un océan à l'autre. Actua met l'accent sur la participation de jeunes sous-représentés dans le cadre de programmes destinés aux Autochtones, aux filles et aux jeunes femmes, aux jeunes à risque ainsi qu'à ceux vivant dans des communautés nordiques ou éloignées. Pour de plus amples renseignements, consultez notre site web à actua.ca et suivez-nous sur [Twitter](#), [Facebook](#), [Instagram](#) et [YouTube](#)!



Annexes

Annexe A – Liens carrières/mentorat

- Ingénieur/ingénieure du son
- Analyste de données
- Analyste en apprentissage automatique
- Programmeur/programmeuse

Annexe B – Information documentaire

Commandes du système de commande des robots (RCS)

Le système s'attend à recevoir les commandes suivantes :

- PORT et STBD (se prononce «starboard») [bâbord et tribord] :
 - PORT actionne le propulseur du côté bâbord, ce qui fait tourner la station dans le sens horaire (vers la droite).
 - STBD actionne le propulseur du côté tribord, ce qui fait tourner la station dans le sens antihoraire (vers la gauche).
- POS (pour « positif») et NEG (pour « négatif») :
 - POS actionne le micropropulseur de tangage positif, ce qui oriente la station et la batterie d'antennes vers le haut.
 - NEG actionne le micropropulseur de tangage négatif, ce qui oriente la station et la batterie d'antennes vers le bas.
- STOP :
 - STOP donne instruction à la station de prendre le contrôle des propulseurs afin de mettre un terme à la manœuvre.

Chaque commande s'écrit TOUT EN MAJUSCULES et s'exécute en une seconde, sauf la commande STOP, qui prend tout le temps nécessaire pour



immobiliser la station. Une fois en mouvement, la station ne s'immobilisera pas toute seule. Un propulseur opposé doit être actionné pour l'immobiliser.

